

### A Sample Information Retrieval Index

Document A : Now is the time.

Document B : For all good men.

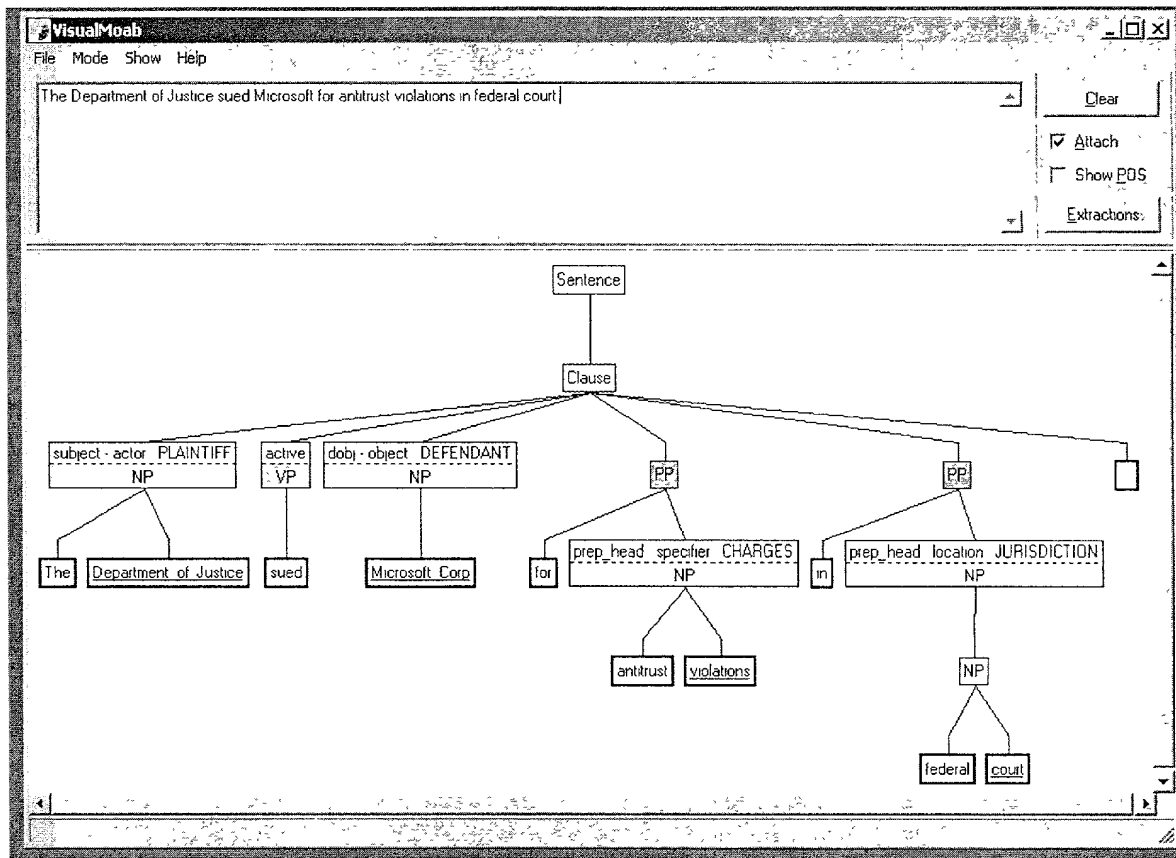
Document C: To come to the aid of their country.

	Document A	Document B	Document C
Now	1		
Is	1		
The	1		1
Time	1		
For		1	
All		1	
Good		1	
Men		1	
To			2
Come			1
Aid			1
of			1
their			1
country			1

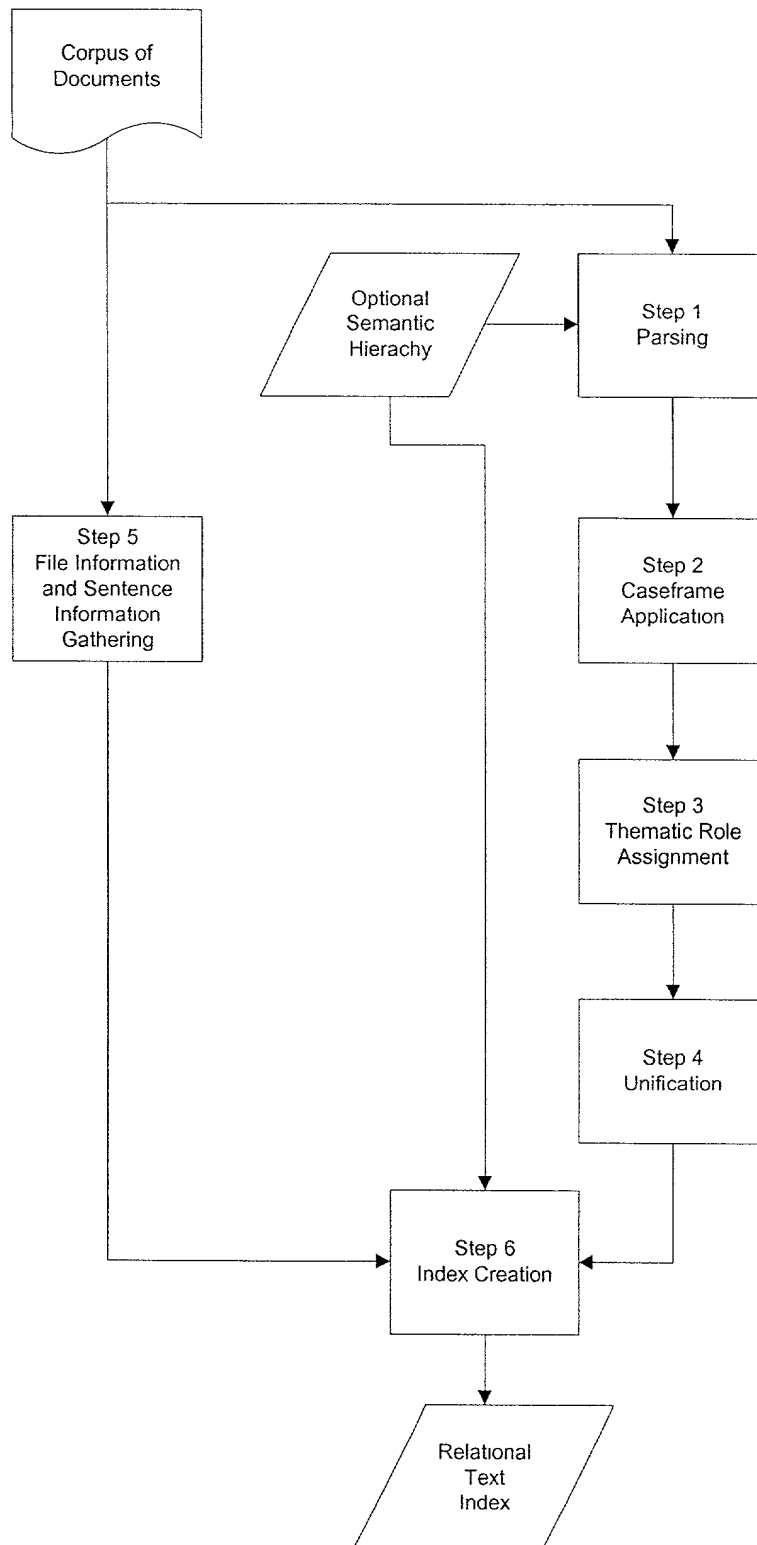
Figure 1

CLAUSE  
NP (SUBJ)  
The [article]  
Department of Justice [noun (phrasal, gov\_entity)]  
VP (ACTIVE\_VOICE)  
sued [verb, past tense]  
NP (DOBJ,)  
Microsoft [noun, (company)]  
PP  
for [preposition]  
NP  
antitrust [adjective]  
violations [noun]  
PP  
in  
NP  
federal [adjective]  
court [noun]

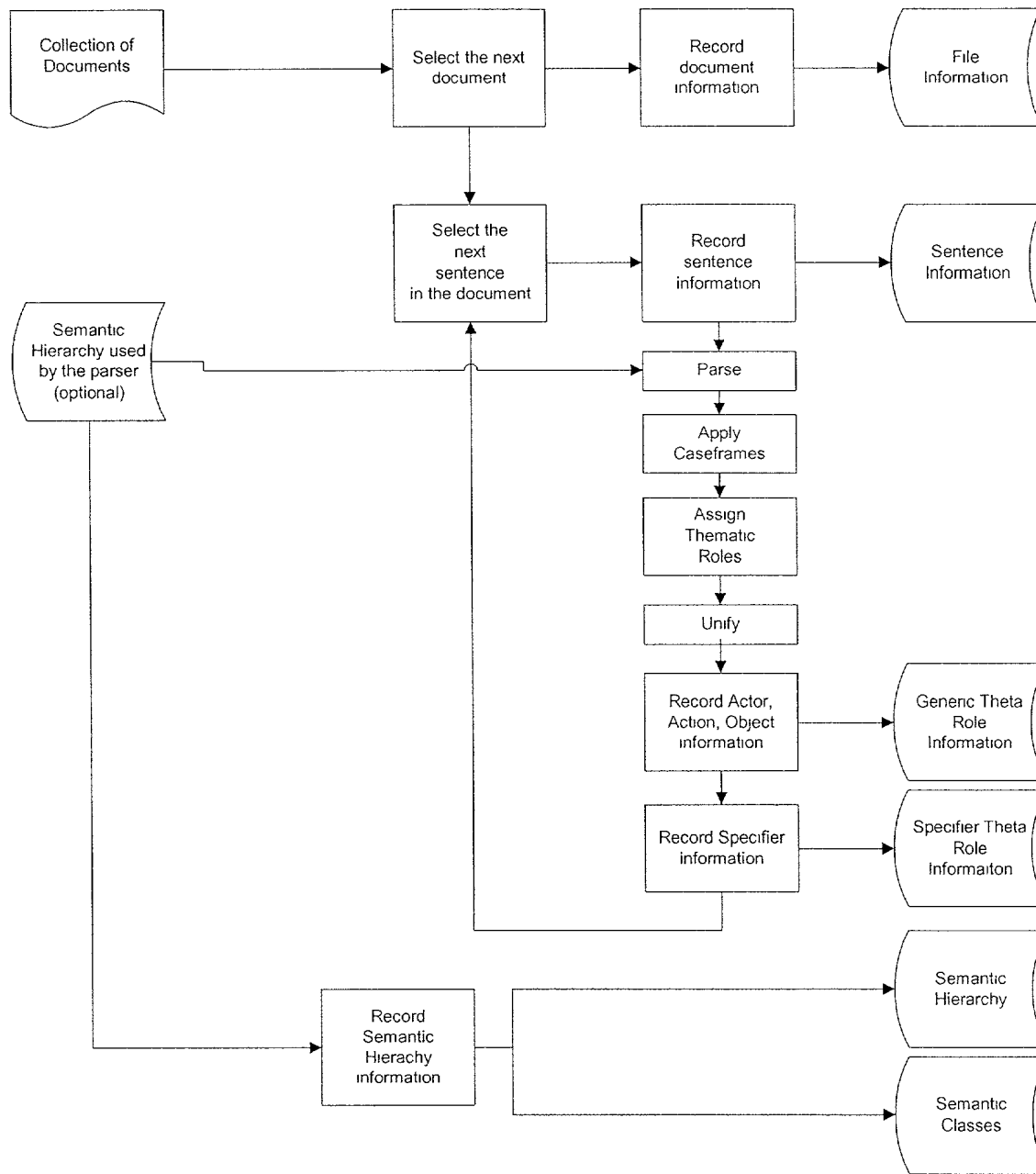
Figure 2a



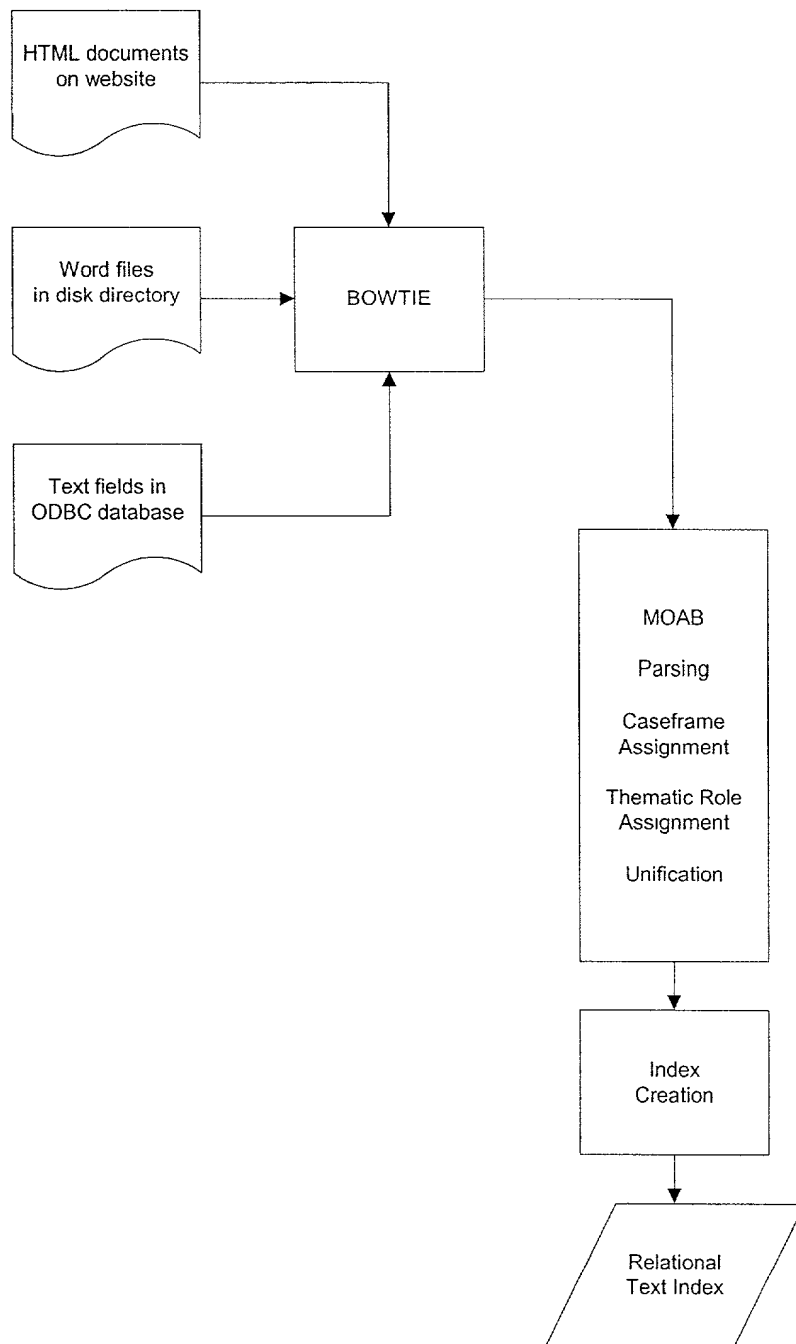
**Figure 2b**  
**A Graphical View of A Sentence Parse and Thematic Role Assignment**



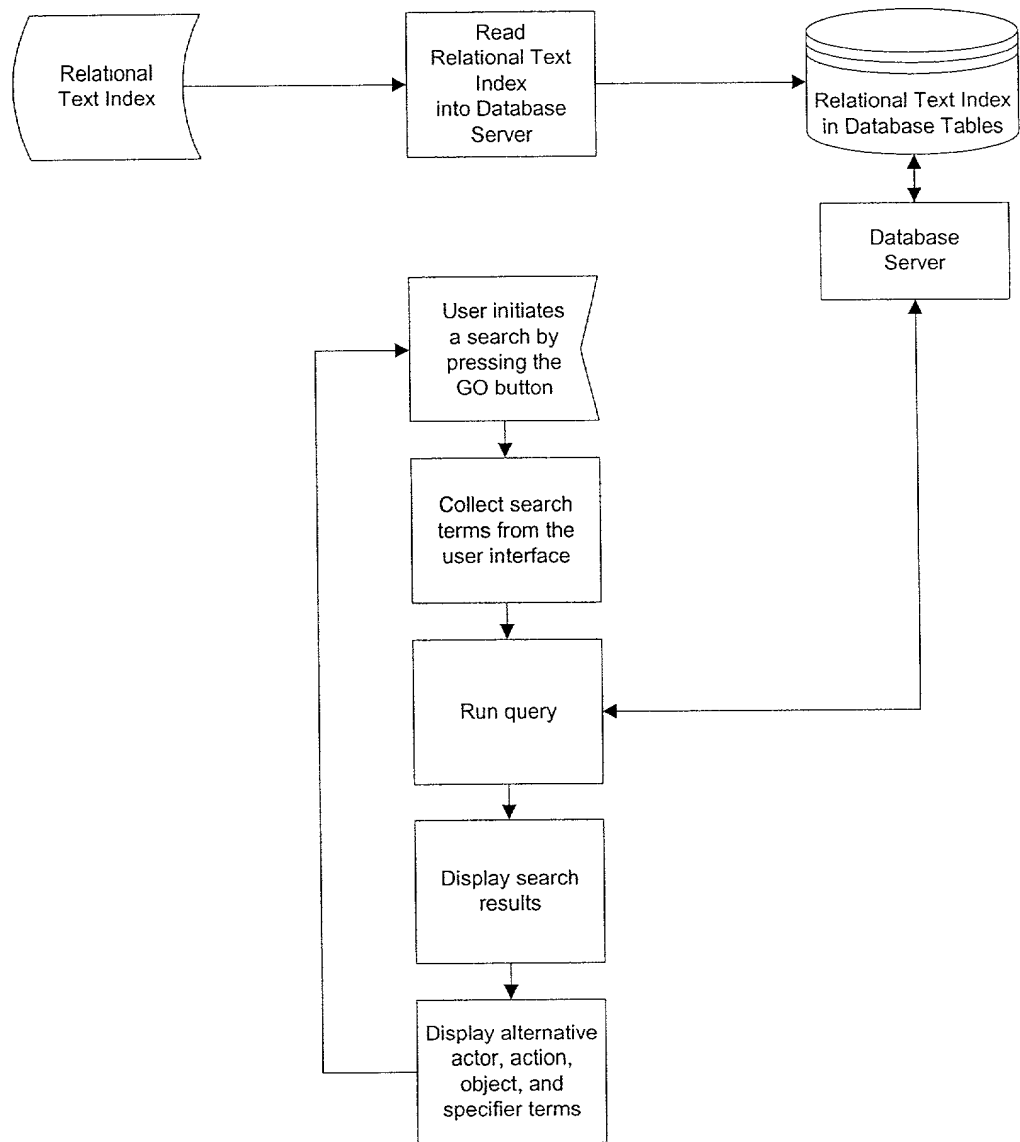
**Figure 3**  
**Index Creation Algorithm, High Level**



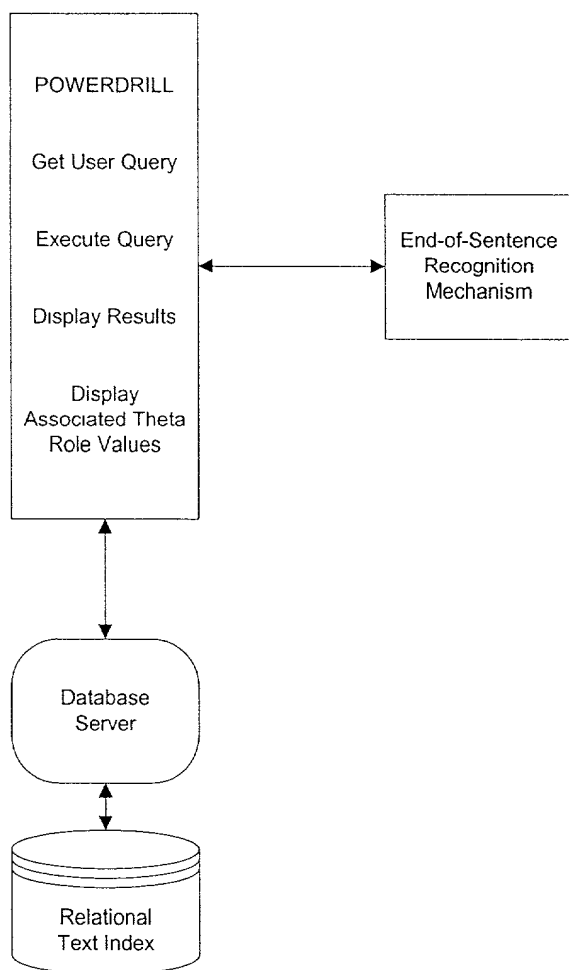
**Figure 4**  
**Index Creation Algorithm, Low Level**



**Figure 5**  
**Overall Flow**



**Figure 6**  
**Flowchart – PowerDrill**



**Figure 7**  
**Overall Search Processing**



**PowerDrill (reuters)** [File] [Wizards] [Favorites] [Settings] [Help]

Integrated | Advanced | What is?

Actor: REAGAN [X] Action: [X] Object: [X] [Clear]

REAGAN

ABIDE  
+ ACCEPT  
ACCOUNTS  
ACCUSED  
+ ACKNOWLEDGE  
ACT  
+ ADDED  
ADDRESS  
+ ADMIT  
+ REAFFIRMED  
+ AGREE  
AID  
ALLEGING  
+ ANNOUNCE  
ANSWERING  
+ APPOINT  
+ APPROVE  
ATTACK  
+ ATTEMPT  
ATTEND  
+ AUTHORIZE  
BE  
BEGIN

1  
1981  
1987  
25  
5-6  
ACID\_RAIN  
+ ACT  
ADDRESS  
ADVICE  
AFTERNOON  
AGENDA  
AGREEMENT  
AID  
ALAN\_GREENSPAN  
ANDEWELT  
ANNIVERSARY  
ANNOUNCEMENT  
ANY  
APPEAL  
APPLAUSE  
ARMS  
AUDIENCE  
AUTHORITY

AND  
OR  
Intersect by  
extraction  
sentence

Results (1-10 of 212) 11:20 >>

Engine	Filename	Filedate	Location
P	reut1003.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10035.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10043.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10127.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10305.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10357.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut1039.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10462.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10
P	reut10735.bt	May 30, 2001	/matrx/tech/corpora/reuters/processed/reut10

Search Completed

FIGURE 8

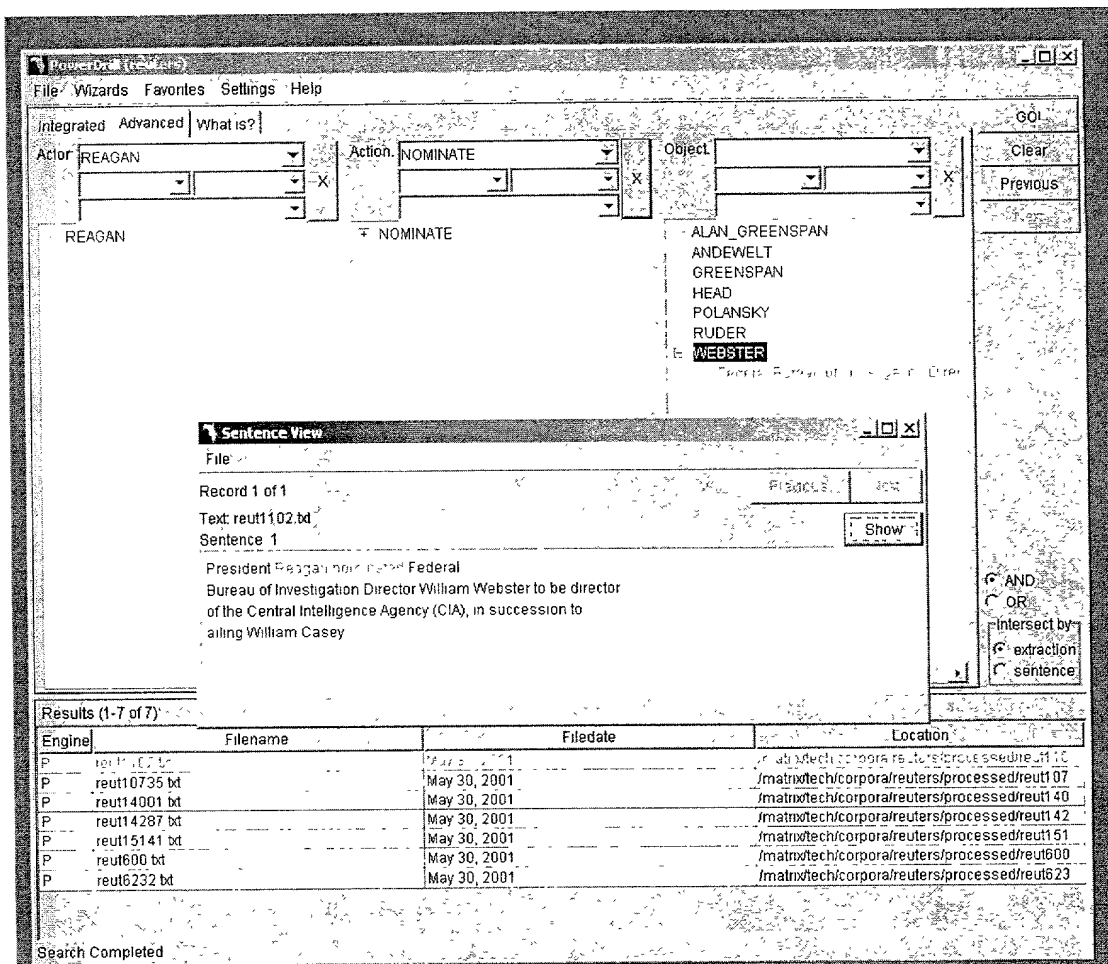


FIGURE 9

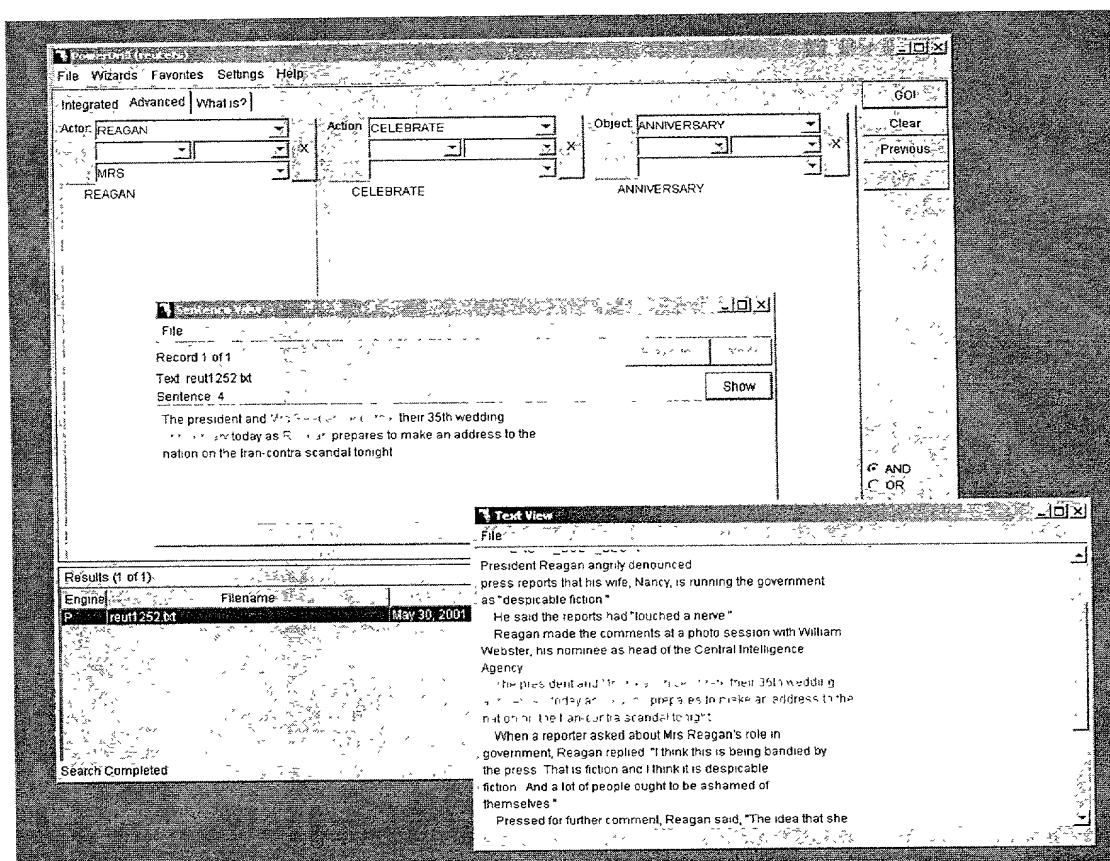


Figure 10

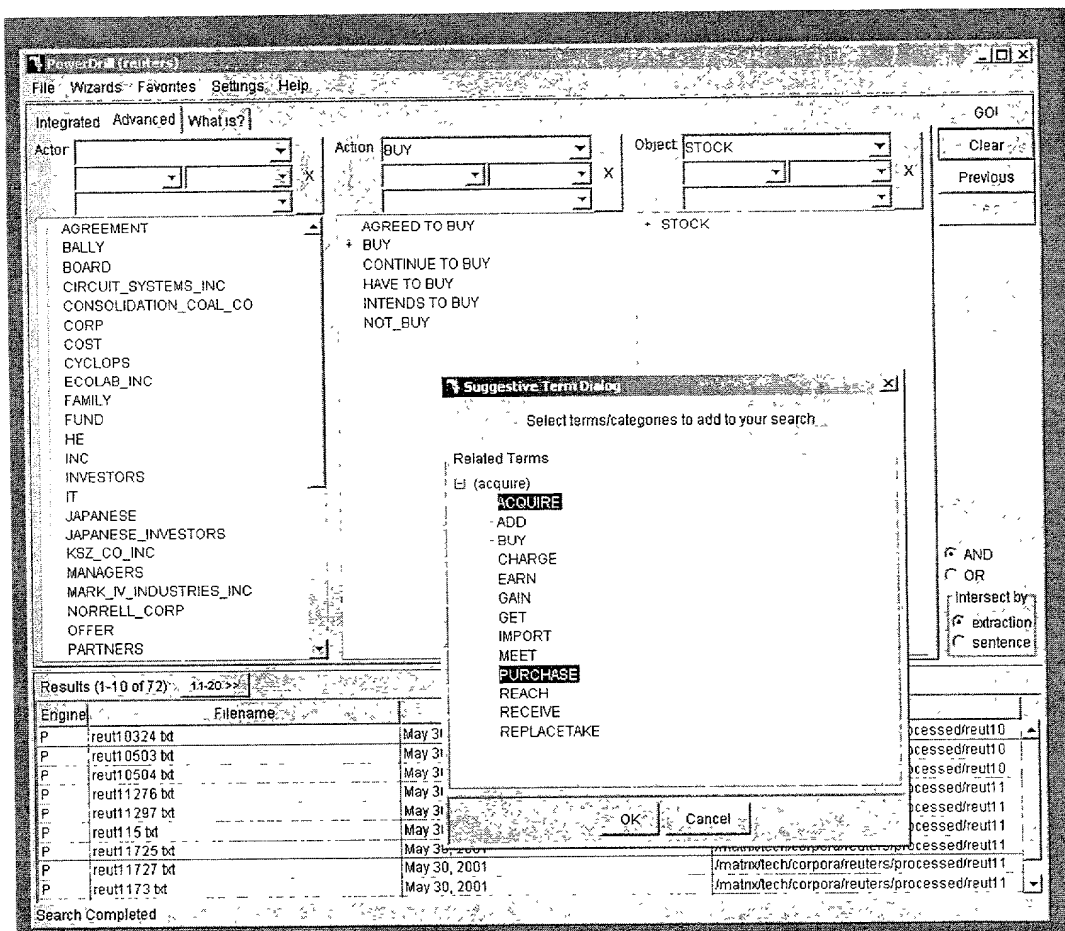


Figure 11

PowerDrill (reuters)

File Wizards Favorites Settings Help

Integrated: Advanced What is?

Actor: INVESTOR

Action: BUY ACQUIRE PURCHASE

Object:

GOI

Clear

Previous

INVESTOR

ACQUIRE

AGREE TO BUY

APPEAR TO BUY

BUY

ENTITLE

FLOCKED TO BUY

PURCHASE

AMOUNT

BENEFITTED

BOND ISSUES

BULLION

CERTIFICATES

COMMON\_SHARE

CONVERTIBLE

DEBT

DOLLARS

FLORIDA

ISSUES

MARK

MEADOWS

METAL

PERCENTAGE

SECURITIES

SHARE

SILVER

STOCKS

THEM

UNITED\_STATES

VESTOR

WARRANT

AND

OR

Intersect by

extraction

sentence

Results (1-10 of 29) 11-20 >>

Engine	Filename	Filedate	Location
P	reut12504 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12542 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12543 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12544 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12545 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12546 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12547 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12548 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12549 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11
P	reut12550 bt	May 30, 2001	/matnotech/corpora/reuters/processed/reut11

Search Completed

Figure 12

**PowerDrill (reuters)** [File] [Wizards] [Favorites] [Settings] [Help]

Integrated Advanced What is?

Actor: [Bank] [BLY] [COMPANY] [DEFENDANTS] [FIRST\_FEDERAL] [HE] [IT] [PERSONS] [RESOLUTION] [SALE] [SHARES] [THRIFT] [TRANS\_WORLD\_AIRLINES\_INC]

Action: VIOLATE

Object: LAW

Clear

Reset

Submit

AND

OR

Intersect by

extraction

sentence

Results (1-10 of 55): 11:20 >>

Engine	Filename	Filedate	Location
P&K	reut7604.txt	May 30, 2001	/matrxtech/corpora/reuters/processed/reut760
The Federal Reserve Board has voted to raise the discount rate to 5.75 percent, the highest level since 1980. The move is seen as a signal that the Fed is concerned about inflation and is prepared to take more aggressive action to curb it.			
P&K	reut4533.txt	May 30, 2001	/matrxtech/corpora/reuters/processed/reut453
The Federal Reserve Board has voted to raise the discount rate to 5.75 percent, the highest level since 1980. The move is seen as a signal that the Fed is concerned about inflation and is prepared to take more aggressive action to curb it.			

Search Completed

Figure 13